

## SEQUENCE LISTING

&lt;110&gt; Diversa Corporation

K. Gray

J. Garrett

N. Aboushadi

R. Knowles

E. O'Donoghue

E. Waters

<120> GLUCOSIDASES, NUCLEIC ACIDS ENCODING  
THEM AND METHODS FOR MAKING AND USING THEM

&lt;130&gt; 564462004040

&lt;150&gt; 60/456,972

&lt;151&gt; 2003-03-20

&lt;160&gt; 24

&lt;170&gt; FastSEQ for Windows Version 4.0

&lt;210&gt; 1

&lt;211&gt; 1710

&lt;212&gt; DNA

&lt;213&gt; Bacteria

&lt;400&gt; 1

atgcagcgcc	atgcagaaag	gtgcagggct	aagcttgctg	ggttcgtttt	ggcaccgccg	60
ttggcgggtg	catggaagcc	cggagggggg	ccctcgatgt	cgcagactcc	atgggtggcg	120
ggtgccgtca	tctaccagat	ctaccgcgcg	agtttcctcg	acgccaacgg	cgacggggtc	180
ggtgacctgc	ccggcatcat	cgaccggctg	gagtacgtgg	ccgcgctggg	cgtggacgcc	240
atctgggtct	cgccgttctt	cacctcgccg	atggccgatt	tcggctacga	catcgccgac	300
catcgcgacg	tggaccggct	gtttggcacg	ctggccgatt	tcgaccggct	gctggccaag	360
gcgcatgcgc	tgggcctgaa	ggtgatgatc	gaccagggtg	tcagccacac	ctcgatcgac	420
cacgcctggt	tccgtgagag	ccggcaggac	cgcaccaatc	cgaaggcgga	ctggtacgtg	480
tgggccgacc	cgcgcgagga	cggcacggcg	cccaacaact	ggatgtcgat	cttcggcggg	540
gtggcctggc	aatgggagcc	gcgcccggag	cagtacttcc	tgcacaactt	cctggccgac	600
cagccggacc	tggatttcca	caaccggcg	gtgcagcagg	ccacgctgga	ctacgtgcgc	660
ttctggctgg	accggggcgt	ggacgggttc	cgcttgact	cgatcaactt	ctgcttccac	720
gacgcgcagt	tgcgcgacaa	cccggccaag	ccgctggaaa	agcgcgtcgg	ccgtggcttc	780
agcgcggaca	atccgtacgc	ctaccagtac	cactactaca	acaacaccca	gccggagaac	840
atcggttca	tcgagcgctt	gcgtgggttg	ctggacgaat	acccgggcac	cgtgagcctg	900
ggcgagatct	cggccgagga	ctcgctggcc	accaccggcg	agtacaccgc	gccgggccgc	960
ctgcacatgg	gctacagctt	cgagctgctg	gtgaaggatt	tcagcgccgg	ctacatccgc	1020
gacaccgtgt	cgcggctgga	agcgacgatg	accgaaggct	ggccgtgctg	ggcgatctcc	1080
aaccacgacg	tggagcgtgc	ggtcactcgc	tggggcgggc	atccggcccg	gccgcggctg	1140
gcgcggatgc	tggtggcgct	gctgtgctcg	ctgcgtggct	cgatctgcct	gtaccagggc	1200
gaggagctgg	gcctgggcga	ggcggacgtg	ccgttcgagg	cgctgcagga	cccgtatggc	1260
atcaccttct	ggccgaactt	caaggggcgc	gacggctgcc	gcacgccgat	gccgtggatc	1320
gatgcgccgt	tggcagggtt	caccagcggt	gagccgtggc	tgccgattcc	ggccgagcac	1380
cgcgccgcgc	cgggtggcgt	gcaggagcac	gaccgcact	cgggtgtgaa	cgcgttccgc	1440
cagttcctgg	catggcgag	gacgatgccg	acgtgtgctg	tgggcgacat	cgtcttcctg	1500
cagacggccg	agccgggtgt	gatgttcgag	cgccggcatg	cgggggagac	gctgtgtgctg	1560
gccttcaacc	tggcggccga	caccgcgcgc	gtggcgctgc	ccgcggcgag	ctggcagccg	1620
atgcacgtgc	cggggccgga	cgtgggcccag	gccgacggcg	ggacgttggt	actgccggcg	1680
cagtcgatgt	actgcgcgcg	cctgggctga				1710

<210> 2  
 <211> 569  
 <212> PRT  
 <213> Bacteria

<220>  
 <221> SIGNAL  
 <222> (1)...(24)

<221> DOMAIN  
 <222> (46)...(434)  
 <223> Alpha amylase, catalytic domain

<400> 2  
 Met Gln Arg His Arg Arg Arg Cys Arg Ala Lys Leu Val Gly Phe Val  
 1 5 10 15  
 Leu Ala Pro Arg Leu Ala Gly Ala Trp Lys Pro Gly Gly Gly Pro Ser  
 20 25 30  
 Met Ser Gln Thr Pro Trp Trp Arg Gly Ala Val Ile Tyr Gln Ile Tyr  
 35 40 45  
 Pro Arg Ser Phe Leu Asp Ala Asn Gly Asp Gly Val Gly Asp Leu Pro  
 50 55 60  
 Gly Ile Ile Asp Arg Leu Glu Tyr Val Ala Ala Leu Gly Val Asp Ala  
 65 70 75 80  
 Ile Trp Val Ser Pro Phe Phe Thr Ser Pro Met Ala Asp Phe Gly Tyr  
 85 90 95  
 Asp Ile Ala Asp His Arg Asp Val Asp Pro Leu Phe Gly Thr Leu Ala  
 100 105 110  
 Asp Phe Asp Arg Leu Leu Ala Lys Ala His Ala Leu Gly Leu Lys Val  
 115 120 125  
 Met Ile Asp Gln Val Phe Ser His Thr Ser Ile Asp His Ala Trp Phe  
 130 135 140  
 Arg Glu Ser Arg Gln Asp Arg Thr Asn Pro Lys Ala Asp Trp Tyr Val  
 145 150 155 160  
 Trp Ala Asp Pro Arg Glu Asp Gly Thr Pro Pro Asn Asn Trp Met Ser  
 165 170 175  
 Ile Phe Gly Gly Val Ala Trp Gln Trp Glu Pro Arg Arg Glu Gln Tyr  
 180 185 190  
 Phe Leu His Asn Phe Leu Ala Asp Gln Pro Asp Leu Asp Phe His Asn  
 195 200 205  
 Pro Ala Val Gln Gln Ala Thr Leu Asp Tyr Val Arg Phe Trp Leu Asp  
 210 215 220  
 Arg Gly Val Asp Gly Phe Arg Leu Asp Ser Ile Asn Phe Cys Phe His  
 225 230 235 240  
 Asp Ala Gln Leu Arg Asp Asn Pro Ala Lys Pro Leu Glu Lys Arg Val  
 245 250 255  
 Gly Arg Gly Phe Ser Ala Asp Asn Pro Tyr Ala Tyr Gln Tyr His Tyr  
 260 265 270  
 Tyr Asn Asn Thr Gln Pro Glu Asn Ile Gly Phe Ile Glu Arg Leu Arg  
 275 280 285  
 Gly Leu Leu Asp Glu Tyr Pro Gly Thr Val Ser Leu Gly Glu Ile Ser  
 290 295 300  
 Ala Glu Asp Ser Leu Ala Thr Thr Ala Glu Tyr Thr Ala Pro Gly Arg  
 305 310 315 320  
 Leu His Met Gly Tyr Ser Phe Glu Leu Leu Val Lys Asp Phe Ser Ala  
 325 330 335  
 Gly Tyr Ile Arg Asp Thr Val Ser Arg Leu Glu Ala Thr Met Thr Glu  
 340 345 350  
 Gly Trp Pro Cys Trp Ala Ile Ser Asn His Asp Val Glu Arg Ala Val

355                      360                      365  
 Thr Arg Trp Gly Gly His Pro Ala Arg Pro Arg Leu Ala Arg Met Leu  
 370                      375                      380  
 Val Ala Leu Leu Cys Ser Leu Arg Gly Ser Ile Cys Leu Tyr Gln Gly  
 385                      390                      395                      400  
 Glu Glu Leu Gly Leu Gly Glu Ala Asp Val Pro Phe Glu Ala Leu Gln  
 405                      410                      415  
 Asp Pro Tyr Gly Ile Thr Phe Trp Pro Asn Phe Lys Gly Arg Asp Gly  
 420                      425                      430  
 Cys Arg Thr Pro Met Pro Trp Ile Asp Ala Pro Leu Ala Gly Phe Thr  
 435                      440                      445  
 Ser Gly Glu Pro Trp Leu Pro Ile Pro Ala Glu His Arg Ala Ala Ala  
 450                      455                      460  
 Val Ala Val Gln Glu His Asp Pro His Ser Val Leu Asn Ala Phe Arg  
 465                      470                      475                      480  
 Gln Phe Leu Ala Trp Arg Arg Thr Met Pro Thr Leu Leu Val Gly Asp  
 485                      490                      495  
 Ile Val Phe Leu Gln Thr Ala Glu Pro Val Leu Met Phe Glu Arg Arg  
 500                      505                      510  
 His Ala Gly Glu Thr Leu Leu Leu Ala Phe Asn Leu Ala Ala Asp Thr  
 515                      520                      525  
 Ala Arg Val Ala Leu Pro Ala Gly Ser Trp Gln Pro Met His Val Pro  
 530                      535                      540  
 Gly Pro Asp Val Gly Gln Ala Asp Gly Gly Thr Leu Val Leu Pro Ala  
 545                      550                      555                      560  
 Gln Ser Met Tyr Cys Ala Arg Leu Gly  
 565

&lt;210&gt; 3

&lt;211&gt; 1293

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;400&gt; 3

atgagtctgt	ggcgtgctgt	cgctcgctgt	gttctgctgt	ccgtcgccat	gcccgcattt	60
gccgacgtgg	tgcgacccgc	ctcttcgcca	ggcgatgtcc	tcaagggtga	gatcaccacc	120
aacggcgagg	gccgcatcgg	ctatgcggtc	acccggctcg	gcaagccggt	aatcggcgag	180
agccacctcg	gattcctcct	ggccgacgcg	ccgcagctgc	tgcgcaactt	ccaggtcgtc	240
gatcaggcca	cccggacctt	cgacgaaacg	tgggagcagc	cgtgggggga	gtggcgacag	300
gtccgcaacc	actacaacga	gctcgcgac	accttcgagg	agaagaccaa	gctccatcgg	360
cggatgcgga	tcgtttttcg	cctgttcgat	gaagggatcg	gcttttcgcta	cgagcttccc	420
cggcagccga	acctggcgca	cgccaacatc	gccgaggaac	tgacccagtt	caacgtcgcg	480
cgaccgggca	cggcctgggtg	ggcaccggcg	ttcgaatcca	accgcgagga	atatctctac	540
aaccagaccc	cgatcgacgg	tgtcgcgatt	gcgatgactc	cattcacgat	gcggttcggag	600
gacgggactc	acctcagcat	ccacgaagcc	gcgctgggtcg	actattccgg	gatgaacgtc	660
acgcgtgtcc	agggcacgaa	cttcaaggcc	atcctgacgc	ccggttcgat	gggccccaaa	720
gtctcccgcg	ataccccat	cgagaccccg	tggcgggtca	tcctgatcag	ccccgacgct	780
gcgcacctct	acgaatcgaa	caggctgac	ctcaacctca	acgaacccaa	caagctcggc	840
gacgtcagct	gggtccaccc	gcgcaaata	gtcgggcat	gggtgggcat	gcacctcgat	900
acccagagtt	gggcctcggg	gccgaagcac	ggcgcgacca	ccgcttatgc	gaagcgaatg	960
atcgatttcg	cggcgacgaa	cggctttacc	gggctgctcg	tcgaaggctg	gaacaaggga	1020
tgggacggag	actggttcgc	gaccggcgac	gatttcagct	tcaccgaacc	ctatcccgat	1080
ttcgacatcc	gagccgtcgc	ggcctacagc	ctcaggaaag	gcgtccacct	catcgggcat	1140
cacgaaacca	gcggcaacat	cgcccactac	gagcaacagc	tcgatgcggc	gctcgatctc	1200
gaccgccagc	tcggcatcga	cacggtgaag	acgggctacg	tctcggatgc	cggcggcac	1260
caggcgctcg	ggcccgcgag	caggatccaa	aga			1293

<210> 4  
 <211> 431  
 <212> PRT  
 <213> Unknown

<220>  
 <223> Obtained from an environmental sample

<221> SIGNAL  
 <222> (1)...(21)

<400> 4  
 Met Ser Leu Trp Arg Ala Leu Val Ala Phe Val Leu Leu Ala Val Ala  
 1 5 10 15  
 Met Pro Ala Phe Ala Asp Val Val Ala Thr Ala Ser Ser Pro Gly Asp  
 20 25 30  
 Val Leu Lys Val Glu Ile Thr Thr Asn Gly Glu Gly Arg Ile Gly Tyr  
 35 40 45  
 Ala Val Thr Arg Leu Gly Lys Pro Val Ile Gly Glu Ser His Leu Gly  
 50 55 60  
 Phe Leu Leu Ala Asp Ala Pro Gln Leu Leu Arg Asn Phe Gln Val Val  
 65 70 75 80  
 Asp Gln Ala Thr Arg Thr Phe Asp Glu Thr Trp Glu Gln Pro Trp Gly  
 85 90 95  
 Glu Trp Arg Thr Val Arg Asn His Tyr Asn Glu Leu Ala Ile Thr Phe  
 100 105 110  
 Glu Glu Lys Thr Lys Leu His Arg Arg Met Arg Ile Val Phe Arg Leu  
 115 120 125  
 Phe Asp Glu Gly Ile Gly Phe Arg Tyr Glu Leu Pro Arg Gln Pro Asn  
 130 135 140  
 Leu Ala His Ala Asn Ile Ala Glu Glu Leu Thr Gln Phe Asn Val Ala  
 145 150 155 160  
 Arg Pro Gly Thr Ala Trp Trp Ala Pro Ala Phe Glu Ser Asn Arg Glu  
 165 170 175  
 Glu Tyr Leu Tyr Asn Gln Thr Pro Ile Asp Gly Val Ala Ile Ala Met  
 180 185 190  
 Thr Pro Phe Thr Met Arg Phe Glu Asp Gly Thr His Leu Ser Ile His  
 195 200 205  
 Glu Ala Ala Leu Val Asp Tyr Ser Gly Met Asn Val Thr Arg Val Gln  
 210 215 220  
 Gly Thr Asn Phe Lys Ala Ile Leu Thr Pro Gly Ser Met Gly Pro Lys  
 225 230 235 240  
 Val Ser Arg Asp Thr Pro Phe Glu Thr Pro Trp Arg Val Ile Leu Ile  
 245 250 255  
 Ser Pro Asp Ala Ala His Leu Tyr Glu Ser Asn Arg Leu Ile Leu Asn  
 260 265 270  
 Leu Asn Glu Pro Asn Lys Leu Gly Asp Val Ser Trp Val His Pro Arg  
 275 280 285  
 Lys Tyr Val Gly Ile Trp Trp Gly Met His Leu Asp Thr Gln Ser Trp  
 290 295 300  
 Ala Ser Gly Pro Lys His Gly Ala Thr Thr Ala Tyr Ala Lys Arg Met  
 305 310 315 320  
 Ile Asp Phe Ala Ala Thr Asn Gly Phe Thr Gly Leu Leu Val Glu Gly  
 325 330 335  
 Trp Asn Lys Gly Trp Asp Gly Asp Trp Phe Ala Thr Gly Asp Asp Phe  
 340 345 350  
 Ser Phe Thr Glu Pro Tyr Pro Asp Phe Asp Ile Arg Ala Val Ala Ala  
 355 360 365

Tyr Ser Leu Arg Lys Gly Val His Leu Ile Gly His His Glu Thr Ser  
 370 375 380  
 Gly Asn Ile Ala His Tyr Glu Gln Gln Leu Asp Ala Ala Leu Asp Leu  
 385 390 395 400  
 Asp Arg Gln Leu Gly Ile Asp Thr Val Lys Thr Gly Tyr Val Ser Asp  
 405 410 415  
 Ala Gly Gly Ile Gln Ala Leu Gly Pro Asp Gly Arg Ile Gln Arg  
 420 425 430

<210> 5  
 <211> 1773  
 <212> DNA  
 <213> Bacteria

<400> 5  
 atgaccctca acaacaccca tgccgactgg tggaaacagg cgggtggtcta ccaggtctac 60  
 ccgcgcagct tcaaggatgt gaacgggtgac ggtctgggcg acatcgccgg cgtgaccgac 120  
 cgcaccccct atctcaagga gctgggcgctc gacgcgatct ggttgctcgcc gttctaccgg 180  
 tccgagctgg cggacggcgg ctacgacgtc atcgactacc gcgacgtgga cccgcgcctg 240  
 ggctccatgg acgacttcga cgccatggcc gcagccgcgc acgaggccgg catgaagggtg 300  
 atcgtggaca tcgtgccgaa ccacacctcc gaccggcacg tctggtttga ggaggcgtg 360  
 gcggcagaac cgggctcccc cgcgcgcgac cggtagatct tccgcgacgg tctgggcgaa 420  
 cacggcgagc tgcccccgaa cgactggcag tcgatcttcg gcggcgggcg atgggaacgg 480  
 gtgcccggac gccagtggta cctgcacatg ttcgccaaagg aacagcccga cctcaactgg 540  
 aagaacccgg aggtccacga ggagttcaag aagaccctgc gtttctggtc ggaccatggc 600  
 gcggacgggt tccgcatcga cgtggcgcac ggcctggcca aggacctcga cagcgcgccg 660  
 ctgaccgaac tggcgttgaa gggcaacccg ctggaggggac tgtgccatga cggcacgaat 720  
 ccgctgtggg accgcccga ggtgcacgac atctaccgcg aatggcgcgga ggtgttcaac 780  
 gactacgac cgccgcgctt cgccgtcgcc gaggcctggg tggtgcccga acaccagcac 840  
 ctgtacgcct ccgaggagga gctggggcag gtgttcaact tcgaattcgc caaggccgac 900  
 tggttcgccg aggatttccg caccgccatc cgcgagggtc tggaaatcagc cgaggaatcc 960  
 ggctccacca ccacctgggt gatgagcaac catgacgtgg tccgccacgc cagccggtac 1020  
 ggctgcccgc aggtgaagac gcagtggtag caccagctgg ccgacgcctg gctgctgcgc 1080  
 agcggttcga acttcatcga ggaccgtgag ctcggcaccc gccgcgctcg cgccgcgac 1140  
 atgatggaac tcggtctgcc cgttccgtg tacgtgtacc agggcgagga gctgggcctg 1200  
 ttcgaggtgg cggacatccc gtgggaccgt ctggaggacc cgaccccgtt ccgcaccgac 1260  
 cacggccgca ccgccaaggg tcgtgacggc tgccgtgtgc cgttgccgtg gtccgccgcc 1320  
 gacgcgccga ccccgccgcc gtgggatgag aagttcggca ccggtgcctc gttcggcttc 1380  
 tccccgtcca tgcgacccga cggcgagccc gccgcgcatc cgcacctgcc gcaaccattg 1440  
 tggtagaagg atttcgcccgc cgatgtggag gactccgac cggattcgat gctcaacctg 1500  
 taccgcaggg ccacgcggtt ccgccagcgg atgtcaccc cgaccggcga tacgtcgac 1560  
 cgggtggatcg gggattccga gttcggcgac aaggccggcc aggtgggtggc gtattcgcg 1620  
 gtcagcacgg acgagggcgg cgaccgcttc gccagcatca ccaacttcgg cccggattcg 1680  
 gtcgagcttc cggccggtga gattgtcttt gcttcgggcc cgctggatgg cgtccggctg 1740  
 cccgccgaca ccaccgtgtg gatgatggtc tga 1773

<210> 6  
 <211> 590  
 <212> PRT  
 <213> Bacteria

<220>  
 <221> DOMAIN  
 <222> (18)...(432)  
 <223> Alpha amylase, catalytic domain

<400> 6  
 Met Thr Leu Asn Asn Thr His Ala Asp Trp Trp Lys Gln Ala Val Val  
 1 5 10 15

Tyr Gln Val Tyr Pro Arg Ser Phe Lys Asp Val Asn Gly Asp Gly Leu  
 20 25 30  
 Gly Asp Ile Ala Gly Val Thr Asp Arg Ile Pro Tyr Leu Lys Glu Leu  
 35 40 45  
 Gly Val Asp Ala Ile Trp Leu Ser Pro Phe Tyr Pro Ser Glu Leu Ala  
 50 55 60  
 Asp Gly Gly Tyr Asp Val Ile Asp Tyr Arg Asp Val Asp Pro Arg Leu  
 65 70 75 80  
 Gly Ser Met Asp Asp Phe Asp Ala Met Ala Ala Ala His Glu Ala  
 85 90 95  
 Gly Met Lys Val Ile Val Asp Ile Val Pro Asn His Thr Ser Asp Arg  
 100 105 110  
 His Val Trp Phe Glu Glu Ala Leu Ala Ala Glu Pro Gly Ser Pro Ala  
 115 120 125  
 Arg Asp Arg Tyr Ile Phe Arg Asp Gly Leu Gly Glu His Gly Glu Leu  
 130 135 140  
 Pro Pro Asn Asp Trp Gln Ser Ile Phe Gly Gly Glu Ala Trp Glu Arg  
 145 150 155 160  
 Val Pro Asp Gly Gln Trp Tyr Leu His Met Phe Ala Lys Glu Gln Pro  
 165 170 175  
 Asp Leu Asn Trp Lys Asn Pro Glu Val His Glu Glu Phe Lys Lys Thr  
 180 185 190  
 Leu Arg Phe Trp Ser Asp His Gly Ala Asp Gly Phe Arg Ile Asp Val  
 195 200 205  
 Ala His Gly Leu Ala Lys Asp Leu Asp Ser Ala Pro Leu Thr Glu Leu  
 210 215 220  
 Ala Leu Lys Gly Asn Pro Leu Glu Gly Leu Cys His Asp Gly Thr Asn  
 225 230 235 240  
 Pro Leu Trp Asp Arg Pro Glu Val His Asp Ile Tyr Arg Glu Trp Arg  
 245 250 255  
 Glu Val Phe Asn Glu Tyr Asp Pro Pro Arg Phe Ala Val Gly Glu Ala  
 260 265 270  
 Trp Val Val Pro Glu His Gln His Leu Tyr Ala Ser Glu Glu Glu Leu  
 275 280 285  
 Gly Gln Val Phe Asn Phe Glu Phe Ala Lys Ala Asp Trp Phe Ala Glu  
 290 295 300  
 Asp Phe Arg Thr Ala Ile Arg Glu Gly Leu Glu Ser Ala Glu Glu Ser  
 305 310 315 320  
 Gly Ser Thr Thr Thr Trp Val Met Ser Asn His Asp Val Val Arg His  
 325 330 335  
 Ala Ser Arg Tyr Gly Leu Pro Gln Val Lys Thr Gln Trp Tyr His Gln  
 340 345 350  
 Leu Ala Asp Ala Trp Leu Leu Arg Ser Gly Ser Asn Phe Ile Glu Asp  
 355 360 365  
 Arg Glu Leu Gly Thr Arg Arg Ala Arg Ala Ala Ile Met Met Glu Leu  
 370 375 380  
 Gly Leu Pro Gly Ser Val Tyr Val Tyr Gln Gly Glu Glu Leu Gly Leu  
 385 390 395 400  
 Phe Glu Val Ala Asp Ile Pro Trp Asp Arg Leu Glu Asp Pro Thr Pro  
 405 410 415  
 Phe Arg Thr Asp His Gly Arg Thr Ala Lys Gly Arg Asp Gly Cys Arg  
 420 425 430  
 Val Pro Leu Pro Trp Ser Ala Ala Asp Ala Pro Thr Pro Ala Pro Trp  
 435 440 445  
 Asp Ala Lys Phe Gly Thr Gly Ala Ser Phe Gly Phe Ser Pro Ser Met  
 450 455 460  
 Arg Pro Asp Gly Glu Pro Ala Ala Asp Pro His Leu Pro Gln Pro Leu  
 465 470 475 480  
 Trp Tyr Lys Asp Phe Ala Ala Asp Val Glu Asp Ser Asp Pro Asp Ser

				485						490					495				
Met	Leu	Asn	Leu	Tyr	Arg	Arg	Ala	Ile	Ala	Phe	Arg	Gln	Arg	Met	Leu				
			500					505					510						
Thr	Pro	Thr	Gly	Asp	Thr	Ser	Ile	Arg	Trp	Ile	Gly	Asp	Ser	Glu	Phe				
		515					520					525							
Gly	Asp	Lys	Ala	Gly	Gln	Val	Val	Ala	Tyr	Ser	Arg	Val	Ser	Thr	Asp				
	530					535					540								
Glu	Gly	Gly	Asp	Arg	Phe	Ala	Ser	Ile	Thr	Asn	Phe	Gly	Pro	Asp	Ser				
545					550					555					560				
Val	Glu	Leu	Pro	Ala	Gly	Glu	Ile	Val	Phe	Ala	Ser	Gly	Pro	Leu	Asp				
			565						570					575					
Gly	Val	Arg	Leu	Pro	Ala	Asp	Thr	Thr	Val	Trp	Met	Met	Val						
		580						585					590						

&lt;210&gt; 7

&lt;211&gt; 2187

&lt;212&gt; DNA

&lt;213&gt; Bacteria

&lt;400&gt; 7

atggtcgggtg	tcgccgcact	cgacgacacc	gtgctacgcg	tagcctactg	cagggtcccct	60
ggggagtggc	cgacgagcac	acctgccatt	gtggaacaga	tgctgcagcg	gcactcctgg	120
cggctcggtt	aggaagagcg	acgcgtgcaa	ctggaatgtg	ttgccggatg	gcagattcaa	180
atcaaccggg	acgacggaac	atggagtata	cgccatctcg	gctttgggac	agccgtcgag	240
gccatcacgt	ggtacaagcg	caaaaaagcg	ggcgcatgga	cgtttgcttc	gctcgacaat	300
gcccggttct	atggacttgg	cgaaaagccg	gggccccttg	acaagcggca	cgaggcgtac	360
accatgtgga	attcggacgt	ttacgcgcct	cacgtgccag	agatggaggc	cctatacctg	420
tcgatcccgt	tctttttgcg	attgcaagac	caaacggccg	ttggcatctt	tgtggataat	480
ccaggggcgat	cccgttttga	cttcaggagc	agataccctg	acgtggaaat	atcgacagaa	540
cgaggcggtt	tggacgttta	tttcatcttc	ggggcttcgc	tcaaagatgt	cattcgacgt	600
tacacaaagc	tcacggggcg	aatgccgatg	ccaccaagt	gggcgctcgg	ctatcatcag	660
tctcgggtact	cgtacgagac	ccagagcgag	gtcctgtcgg	tggcgcaaac	gtttgtcgag	720
cgcgacattc	cgggtggatgc	cttgatatctg	gatattcact	acatggatgg	atatcggttg	780
ttcacctttg	atgagaggcg	atttcccgat	ccggcacgca	tgtgcgacga	gttgcggaag	840
cttgggggtg	gcgttgtccc	catcgtggat	ccgggggtga	aacaggatcc	tgaatatccc	900
gtttacatgg	acgggcttagc	ccacaaccac	ttctgccaaa	ccgccgaggg	tcaagtgtat	960
cttggtgagg	tatggcccgg	attgagtgcc	tttcttgatt	tcgcgtcgga	ggaagtgcgg	1020
gcgtggtggg	gcaagtggca	tcgggtgtac	acccagatgg	gaattgaggg	catatggaat	1080
gatatgaacg	agccggctgt	gttcaatgag	acgaagacca	tggacgtgaa	cgtgggtccac	1140
cgaggcgacg	gacgattgta	cacgcacgga	gaggtccaca	acctctatgg	attctggatg	1200
gctgaagcca	cgtatcgcg	actgaaagcg	cagttggctg	gcaagcggcc	ttttgtgctc	1260
acgcgcgctg	gctacagcgg	aattcaacgg	tatgcggcgg	tgtggacggg	cgacaaccga	1320
agctttttgg	aacatatggc	catggccatc	ccgatgggtg	tcaacatggg	gatgtccggt	1380
attccgctcg	ggggcccaga	tgtgggcgg	ttcgtctatc	acgcttcagg	tgaattgctt	1440
gcacgatgga	cgaaaatggg	ggcggttctt	ccgttctttc	gaaatcacag	tgccatggga	1500
acgcacgcac	aggaaccttg	ggcattcggc	cctacgttcg	aagccgtcat	tcgacgtgcc	1560
atccggctcc	gttatcgttt	cctgccttat	ctgtacacgc	tcgcacggga	agcccatgag	1620
acagggtttac	ccatgatgcg	gcccttggtg	ttagagtatc	ccgacgatcc	aaacacgcac	1680
cacgtcgatg	atcagttcct	cgtcggttcc	gatcttctcg	tggcacccat	cctcaagccg	1740
ggcatggccc	atagaatgg	ctatttgccc	gatggagaat	ggattgatta	tgagacgcgt	1800
gagcgatacc	aagggcgctca	gtacatcttg	acgtatcgc	ctctcgatcg	tatccctctg	1860
tacgtgaggg	cagggagcgc	tatccccgtg	aacctcttgg	agcggtcggg	cgagaccag	1920
ctaggatggg	agatcttcgt	ggacgccaac	ggtcgagcct	cgggtcgatg	ctatgaggac	1980
gacggcgaga	cgtttagcta	tgaagacggc	gcctactgtg	atcgcgtgtt	acaggccctc	2040
gccacctcgg	aaggaacact	gatcgaatgt	catcttggtc	aaggatcagg	agatgggtgga	2100
agtctcgaga	gcgttggttcg	cgtgttcaca	ccagatgacg	ttcgtgaggc	tcgtgcgcag	2160
ggcatatcgt	tttccatcca	tgtgtga				2187

&lt;210&gt; 8

<211> 728  
 <212> PRT  
 <213> Bacteria

<220>  
 <221> DOMAIN  
 <222> (58)...(717)  
 <223> Glycosyl hydrolases family 31

<400> 8  
 Met Val Gly Val Ala Ala Leu Asp Asp Thr Val Leu Arg Val Ala Tyr  
 1 5 10 15  
 Cys Arg Ser Pro Gly Glu Trp Pro Thr Ser Thr Pro Ala Ile Val Glu  
 20 25 30  
 Gln Met Ser Gln Arg His Ser Trp Arg Leu Val Gln Glu Arg Arg  
 35 40 45  
 Val Gln Leu Glu Cys Val Ala Gly Trp Gln Ile Gln Ile Asn Arg Asp  
 50 55 60  
 Asp Gly Thr Trp Ser Ile Arg His Leu Gly Phe Gly Thr Ala Val Glu  
 65 70 75 80  
 Ala Ile Thr Trp Tyr Lys Arg Lys Lys Gly Gly Ala Leu Thr Phe Ala  
 85 90 95  
 Ser Leu Asp Asn Ala Arg Phe Tyr Gly Leu Gly Glu Lys Pro Gly Pro  
 100 105 110  
 Leu Asp Lys Arg His Glu Ala Tyr Thr Met Trp Asn Ser Asp Val Tyr  
 115 120 125  
 Ala Pro His Val Pro Glu Met Glu Ala Leu Tyr Leu Ser Ile Pro Phe  
 130 135 140  
 Phe Leu Arg Leu Gln Asp Gln Thr Ala Val Gly Ile Phe Val Asp Asn  
 145 150 155 160  
 Pro Gly Arg Ser Arg Phe Asp Phe Arg Ser Arg Tyr Pro Asp Val Glu  
 165 170 175  
 Ile Ser Thr Glu Arg Gly Gly Leu Asp Val Tyr Phe Ile Phe Gly Ala  
 180 185 190  
 Ser Leu Lys Asp Val Ile Arg Arg Tyr Thr Lys Leu Thr Gly Arg Met  
 195 200 205  
 Pro Met Pro Pro Lys Trp Ala Leu Gly Tyr His Gln Ser Arg Tyr Ser  
 210 215 220  
 Tyr Glu Thr Gln Ser Glu Val Leu Ser Val Ala Gln Thr Phe Val Glu  
 225 230 235 240  
 Arg Asp Ile Pro Val Asp Ala Leu Tyr Leu Asp Ile His Tyr Met Asp  
 245 250 255  
 Gly Tyr Arg Val Phe Thr Phe Asp Glu Arg Arg Phe Pro Asp Pro Ala  
 260 265 270  
 Arg Met Cys Asp Glu Leu Arg Lys Leu Gly Val Arg Val Val Pro Ile  
 275 280 285  
 Val Asp Pro Gly Val Lys Gln Asp Pro Glu Tyr Pro Val Tyr Met Asp  
 290 295 300  
 Gly Leu Ala His Asn His Phe Cys Gln Thr Ala Glu Gly Gln Val Tyr  
 305 310 315 320  
 Leu Gly Glu Val Trp Pro Gly Leu Ser Ala Phe Pro Asp Phe Ala Ser  
 325 330 335  
 Glu Glu Val Arg Ala Trp Trp Gly Lys Trp His Arg Val Tyr Thr Gln  
 340 345 350  
 Met Gly Ile Glu Gly Ile Trp Asn Asp Met Asn Glu Pro Ala Val Phe  
 355 360 365  
 Asn Glu Thr Lys Thr Met Asp Val Asn Val Val His Arg Gly Asp Gly  
 370 375 380  
 Arg Leu Tyr Thr His Gly Glu Val His Asn Leu Tyr Gly Phe Trp Met



385                      390                      395                      400  
 Ala Glu Ala Thr Tyr Arg Gly Leu Lys Ala Gln Leu Ala Gly Lys Arg  
                                  405                      410                      415  
 Pro Phe Val Leu Thr Arg Ala Gly Tyr Ser Gly Ile Gln Arg Tyr Ala  
                                  420                      425                      430  
 Ala Val Trp Thr Gly Asp Asn Arg Ser Phe Trp Glu His Met Ala Met  
                                  435                      440                      445  
 Ala Ile Pro Met Val Leu Asn Met Gly Met Ser Gly Ile Pro Leu Gly  
                                  450                      455                      460  
 Gly Pro Asp Val Gly Gly Phe Ala His His Ala Ser Gly Glu Leu Leu  
                                  465                      470                      475                      480  
 Ala Arg Trp Thr Gln Met Gly Ala Phe Phe Pro Phe Phe Arg Asn His  
                                  485                      490                      495  
 Ser Ala Met Gly Thr His Arg Gln Glu Pro Trp Ala Phe Gly Pro Thr  
                                  500                      505                      510  
 Phe Glu Ala Val Ile Arg Arg Ala Ile Arg Leu Arg Tyr Arg Phe Leu  
                                  515                      520                      525  
 Pro Tyr Leu Tyr Thr Leu Ala Arg Glu Ala His Glu Thr Gly Leu Pro  
                                  530                      535                      540  
 Met Met Arg Pro Leu Val Leu Glu Tyr Pro Asp Asp Pro Asn Thr His  
                                  545                      550                      555                      560  
 His Val Asp Asp Gln Phe Leu Val Gly Ser Asp Leu Leu Val Ala Pro  
                                  565                      570                      575  
 Ile Leu Lys Pro Gly Met Ala His Arg Met Val Tyr Leu Pro Asp Gly  
                                  580                      585                      590  
 Glu Trp Ile Asp Tyr Glu Thr Arg Glu Arg Tyr Gln Gly Arg Gln Tyr  
                                  595                      600                      605  
 Ile Leu Thr Tyr Ala Pro Leu Asp Arg Ile Pro Leu Tyr Val Arg Ala  
                                  610                      615                      620  
 Gly Ser Ala Ile Pro Val Asn Leu Leu Glu Arg Ser Gly Glu Thr Gln  
                                  625                      630                      635                      640  
 Leu Gly Trp Glu Ile Phe Val Asp Ala Asn Gly Arg Ala Ser Gly Arg  
                                  645                      650                      655  
 Cys Tyr Glu Asp Asp Gly Glu Thr Phe Ser Tyr Glu Asp Gly Ala Tyr  
                                  660                      665                      670  
 Cys Asp Arg Val Leu Gln Ala Leu Ala Thr Ser Glu Gly Thr Leu Ile  
                                  675                      680                      685  
 Glu Cys His Leu Val Gln Gly Ser Gly Asp Gly Gly Ser Leu Glu Ser  
                                  690                      695                      700  
 Val Val Arg Val Phe Thr Pro Asp Asp Val Arg Glu Ala Arg Ala Gln  
                                  705                      710                      715                      720  
 Gly Ile Ser Phe Ser Ile His Val  
                                  725

&lt;210&gt; 9

&lt;211&gt; 1611

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;400&gt; 9

atgacagagt	ggtggcgtgg	tgcaagtgacc	tatcaagtct	atccaaggctc	gtttcaggac	60
agcaacggcg	acggcatcgg	cgacctgccc	ggcatcaccg	cccggcttga	gtatctggcc	120
gatcttggcg	tggacgcgg	ctggctgtca	ccgttcttca	aaagcccgat	gaaggacatg	180
ggctatgacg	tcagcgacta	ttgcgatgtc	gatccggtct	tcggcaccct	cgccgatttt	240
gacgccctgc	tggcccgcgc	gcatgagctg	gggctcaagg	tgatcatcga	ccaggtccct	300
agccacagtt	ccgacctgca	ccctgccttt	gtgaccagtc	gcagcgaccg	cgtgaaccgc	360

```

aaggcggact ggtatgtctg ggccgatccc aagcccgacg gcagcccgcc caacaactgg 420
ctgtcgggtg tccgtggctc ggcatgggccc tgggacgcgc gcagaaaaca gtattacctg 480
cacaatttcc tgaccagcca gccggacctg aactaccaca acccgaaggt gcaggactgg 540
gcgctggaca acatgcgttt ctggctggac cggggcgctgg acgggttccg ctttgacacc 600
gtcaactact tcttccacga tcccttggtg cgcagcaacc ctgccgatca ccgcaacaag 660
cctgaggctg acggcaatcc ctacggcatg cagtaccacc tgcatgacaa gaaccagccc 720
gagaacctga tctggatgga gcggatacgg gtgcttctgg accaatacgg tgccgcaagc 780
gtcggcgaga tgggcgaaaag tcaccaacgc atccggatga tgggcgacta caccgctccg 840
gggaggctgc atcaatgcta cagctttgaa ttcattgggt atgaatacac cgcaaacctg 900
ttccgggacc ggatagaaag ctttttcaag ggtgccccta aaggctggcc gatgtggcg 960
ttttcaaacc acgatgtcgt ccgccatgtc agtcgctggg caaaacatgg cctcaccccc 1020
gaggcgggtt ccaagcagac aggtgcgttg cttctgtcgc ttgagggtc gatctgcctg 1080
tgggagggcg aggagctggg ccagaccgat accgaactgg ccttgatga gttgaccgat 1140
ccgcagggca tcgtcttttg gcccgaaacc atcggccgcg acaatactcg gacgccaatg 1200
gtttgggacg catcgccgca tggcgggttt tcgaccgtca caccctggct gccggtgaaa 1260
ccggaacagg ccgcgcgtca tgtggccggg caaacccgtg atgccgcctc ggtgctggaa 1320
agctaccggg cgatgctggc cttccggcgc gctgaaccgg cccttaggac cgggcggacg 1380
cggtttcttg atctggccga accggttctg ggctttgtgc gcggcgaagg ggagggtcg 1440
atcctgtgcc tgttcaatct gtcgcctgtt gcgcgggggg ttgcggtcga aggcgtgggc 1500
ccgccgatcg gcccgggcca gcaggctatc ctttcgggcg gacggctagg ccttgggccc 1560
aacggcgccg ctttctctgc ggtgaccgga acagtccgcg ttctggacta a 1611

```

&lt;210&gt; 10

&lt;211&gt; 536

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;221&gt; DOMAIN

&lt;222&gt; (12)...(449)

&lt;223&gt; Alpha amylase, catalytic domain

&lt;400&gt; 10

```

Met Thr Glu Trp Trp Arg Gly Ala Val Thr Tyr Gln Val Tyr Pro Arg
 1          5          10          15
Ser Phe Gln Asp Ser Asn Gly Asp Gly Ile Gly Asp Leu Pro Gly Ile
 20          25          30
Thr Ala Arg Leu Glu Tyr Leu Ala Asp Leu Gly Val Asp Ala Val Trp
 35          40          45
Leu Ser Pro Phe Phe Lys Ser Pro Met Lys Asp Met Gly Tyr Asp Val
 50          55          60
Ser Asp Tyr Cys Asp Val Asp Pro Val Phe Gly Thr Leu Ala Asp Phe
 65          70          75          80
Asp Ala Leu Leu Ala Arg Ala His Glu Leu Gly Leu Lys Val Ile Ile
 85          90          95
Asp Gln Val Leu Ser His Ser Ser Asp Leu His Pro Ala Phe Val Thr
100          105          110
Ser Arg Ser Asp Arg Val Asn Pro Lys Ala Asp Trp Tyr Val Trp Ala
115          120          125
Asp Pro Lys Pro Asp Gly Ser Pro Pro Asn Asn Trp Leu Ser Val Phe
130          135          140
Gly Gly Ser Ala Trp Ala Trp Asp Ala Arg Arg Lys Gln Tyr Tyr Leu
145          150          155          160
His Asn Phe Leu Thr Ser Gln Pro Asp Leu Asn Tyr His Asn Pro Lys
165          170          175
Val Gln Asp Trp Ala Leu Asp Asn Met Arg Phe Trp Leu Asp Arg Gly
180          185          190

```

Val Asp Gly Phe Arg Phe Asp Thr Val Asn Tyr Phe Phe His Asp Pro  
 195 200 205  
 Leu Leu Arg Ser Asn Pro Ala Asp His Arg Asn Lys Pro Glu Ala Asp  
 210 215 220  
 Gly Asn Pro Tyr Gly Met Gln Tyr His Leu His Asp Lys Asn Gln Pro  
 225 230 235 240  
 Glu Asn Leu Ile Trp Met Glu Arg Ile Arg Val Leu Leu Asp Gln Tyr  
 245 250 255  
 Gly Ala Ala Ser Val Gly Glu Met Gly Glu Ser His His Ala Ile Arg  
 260 265 270  
 Met Met Gly Asp Tyr Thr Ala Pro Gly Arg Leu His Gln Cys Tyr Ser  
 275 280 285  
 Phe Glu Phe Met Gly Tyr Glu Tyr Thr Ala Asn Leu Phe Arg Asp Arg  
 290 295 300  
 Ile Glu Ser Phe Phe Lys Gly Ala Pro Lys Gly Trp Pro Met Trp Ala  
 305 310 315 320  
 Phe Ser Asn His Asp Val Val Arg His Val Ser Arg Trp Ala Lys His  
 325 330 335  
 Gly Leu Thr Pro Glu Ala Val Ala Lys Gln Thr Gly Ala Leu Leu Leu  
 340 345 350  
 Ser Leu Glu Gly Ser Ile Cys Leu Trp Glu Gly Glu Glu Leu Gly Gln  
 355 360 365  
 Thr Asp Thr Glu Leu Ala Leu Asp Glu Leu Thr Asp Pro Gln Gly Ile  
 370 375 380  
 Val Phe Trp Pro Glu Pro Ile Gly Arg Asp Asn Thr Arg Thr Pro Met  
 385 390 395 400  
 Val Trp Asp Ala Ser Pro His Gly Gly Phe Ser Thr Val Thr Pro Trp  
 405 410 415  
 Leu Pro Val Lys Pro Glu Gln Ala Ala Arg His Val Ala Gly Gln Thr  
 420 425 430  
 Gly Asp Ala Ala Ser Val Leu Glu Ser Tyr Arg Ala Met Leu Ala Phe  
 435 440 445  
 Arg Arg Ala Glu Pro Ala Leu Arg Thr Gly Arg Thr Arg Phe Leu Asp  
 450 455 460  
 Leu Ala Glu Pro Val Leu Gly Phe Val Arg Gly Glu Gly Glu Gly Ala  
 465 470 475 480  
 Ile Leu Cys Leu Phe Asn Leu Ser Pro Val Ala Arg Gly Val Ala Val  
 485 490 495  
 Glu Gly Val Gly Pro Pro Ile Gly Pro Gly Gln Gln Ala Ile Leu Ser  
 500 505 510  
 Gly Gly Arg Leu Gly Leu Gly Pro Asn Gly Ala Ala Phe Leu Arg Val  
 515 520 525  
 Thr Gly Thr Val Arg Val Leu Asp  
 530 535

&lt;210&gt; 11

&lt;211&gt; 1719

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt; .

&lt;223&gt; Obtained from an environmental sample

&lt;400&gt; 11

gtgcccgaac	tgaaatggtg	gcagacagct	atattttatc	aaatctaccc	gcgctctttc	60
gccgacggca	acggcgatgg	catcgggcat	ttcaaaggca	tcacgaggca	actcgattat	120
ctacaaaatc	ttggcataga	tgcgctctgg	ctctcgccctc	acttcccctc	ccccaaactgg	180
gattgcggtc	acgatatcag	cgattaccgc	aacgttgctc	cggaatacgg	cacgctggac	240
gatttcaaaa	ccttcctgag	cgaatcgcac	aaacgcggta	tccgcgtcat	tctcgacctc	300

```

gtgctgaatc acacctccga tgaacatccg tggttcatcg aatcgaaatc cagccgcgat 360
aatcccaa atccgattggta tgtgtgggtc gatacgccgc ccaacaattg gcagtcctgc 420
ttcgatggcg atgcctggac atacgtccct gaacgcggcc aatattatta tcaactacttc 480
atgaaacagc agcccgatct caactggcat aatccgcagg tcaaacaggc catgtgggag 540
gcggtgcgct tctggctcga tctcggcgtg gacgggttcc gcctggacgc catcggcacg 600
atctacgaag acccaaatct cacgccgcat aatgtcccga tgaatttggc tgagctgcgt 660
cacttcacag atgtcgccaa aacgccggaa gagatcaagc tcaaagaaaa atactggcac 720
gacatgttca agcatcaatg gggtcagccc ggcgttcacg acctgatgaa agaactgcgc 780
gccatcctcg atgaatatga tggcgaccgc atgtctgtcg gcgaagatga caacatcgat 840
tacatgggca acggagacga cgaattgcag ctggtcttca acttcccgtt gatgcgcgcc 900
gatcgtctca cccccgacca tattcggcgc aacaaaaaag agcgtttgac tcgtctgaat 960
gctttaccgg ttaaaggctg ggcttgcaac acgctcggca accatgatag ttcacgcgtc 1020
tacaccaaat tcggtgacg gatccacggc gcggaccatg cacgtctcaa cctggcgctt 1080
ttgctcaccg tgcacggcac gccgttctta tacaacggcg aagagatcgg catgaccgac 1140
cacatcatta ccgatcccac caaactgcgc gagcccgccc ttcgcgcggg ttacaacagc 1200
cttgtcaacg aaatgaaggt cgagccagcg gagggccccc ttcgcgcggg acagatgacg 1260
cgcgacaaaa accgtacccc catgcaatgg gacaataagc ccaatgccgg tttttgccc 1320
gataaagccg aaccctggtt gccagtcaac cccaattacc gcgcaggcat taacgtccgc 1380
gagcaaacat cgaacccgaa ctcgctgctc aattactata aacgtctcat ccacttgccg 1440
cgggaaacgc ctgccttgat cgctggagat tacgttccgc ttcaccagac atccaaagat 1500
catctggcct tcctgcgcaa aacagattca caacgatcc tggtcgtttt gaattactcc 1560
cccaataaat tggaattgga tttctcgcgc accgtcgaaa tgaaaggccg cccgctgatc 1620
gcaattttct ccagcgcaga tgaccgcccg caggcggcac aaagcccaaa gaaagtatcg 1680
gtcggcgctt acggagtctt gctggcagaa gtaaaatag 1719

```

&lt;210&gt; 12

&lt;211&gt; 572

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;221&gt; DOMAIN

&lt;222&gt; (14)...(480)

&lt;223&gt; Alpha amylase, catalytic domain

&lt;400&gt; 12

```

Met Pro Glu Leu Lys Trp Trp Gln Thr Ala Ile Phe Tyr Gln Ile Tyr
1          5          10          15
Pro Arg Ser Phe Ala Asp Gly Asn Gly Asp Gly Ile Gly Asp Phe Lys
20          25          30
Gly Ile Ile Gly Lys Leu Asp Tyr Leu Gln Asn Leu Gly Ile Asp Ala
35          40          45
Leu Trp Leu Ser Pro His Phe Pro Ser Pro Asn Trp Asp Cys Gly Tyr
50          55          60
Asp Ile Ser Asp Tyr Arg Asn Val Ala Pro Glu Tyr Gly Thr Leu Asp
65          70          75          80
Asp Phe Lys Thr Phe Leu Ser Glu Ser His Lys Arg Gly Ile Arg Val
85          90          95
Ile Leu Asp Leu Val Leu Asn His Thr Ser Asp Glu His Pro Trp Phe
100          105          110
Ile Glu Ser Lys Ser Ser Arg Asp Asn Pro Lys Ser Asp Trp Tyr Val
115          120          125
Trp Val Asp Thr Pro Pro Asn Asn Trp Gln Ser Cys Phe Asp Gly Asp
130          135          140
Ala Trp Thr Tyr Val Pro Glu Arg Gly Gln Tyr Tyr Tyr His Tyr Phe
145          150          155          160
Met Lys Gln Gln Pro Asp Leu Asn Trp His Asn Pro Gln Val Lys Gln

```

165 170 175  
 Ala Met Trp Glu Ala Val Arg Phe Trp Leu Asp Leu Gly Val Asp Gly  
 180 185 190  
 Phe Arg Leu Asp Ala Ile Gly Thr Ile Tyr Glu Asp Pro Asn Leu Thr  
 195 200 205  
 Pro His Asn Val Pro Met Asn Leu Ala Glu Leu Arg His Phe Thr Asp  
 210 215 220  
 Val Ala Lys Thr Pro Glu Glu Ile Lys Leu Lys Glu Lys Tyr Trp His  
 225 230 235 240  
 Asp Met Phe Lys His Gln Trp Gly Gln Pro Gly Val His Asp Leu Met  
 245 250 255  
 Lys Glu Leu Arg Ala Ile Leu Asp Glu Tyr Asp Gly Asp Arg Met Leu  
 260 265 270  
 Val Gly Glu Asp Asp Asn Ile Asp Tyr Met Gly Asn Gly Asp Asp Glu  
 275 280 285  
 Leu Gln Leu Val Phe Asn Phe Pro Leu Met Arg Ala Asp Arg Leu Thr  
 290 295 300  
 Pro Asp His Ile Arg Arg Asn Gln Lys Glu Arg Leu Thr Arg Leu Asn  
 305 310 315 320  
 Ala Leu Pro Val Lys Gly Trp Ala Cys Asn Thr Leu Gly Asn His Asp  
 325 330 335  
 Ser Ser Arg Val Tyr Thr Lys Phe Gly Asp Arg Ile His Gly Ala Asp  
 340 345 350  
 His Ala Arg Leu Asn Leu Ala Leu Leu Leu Thr Leu His Gly Thr Pro  
 355 360 365  
 Phe Leu Tyr Asn Gly Glu Glu Ile Gly Met Thr Asp His Ile Ile Thr  
 370 375 380  
 Asp Pro Thr Lys Leu Arg Asp Thr Met Ala Thr Trp Tyr Tyr Asn Ser  
 385 390 395 400  
 Leu Val Asn Glu Met Lys Val Glu Pro Ala Glu Ala Ala Leu Arg Ala  
 405 410 415  
 Gly Gln Met Thr Arg Asp Lys Asn Arg Thr Pro Met Gln Trp Asp Asn  
 420 425 430  
 Lys Pro Asn Ala Gly Phe Cys Pro Asp Lys Ala Glu Pro Trp Leu Pro  
 435 440 445  
 Val Asn Pro Asn Tyr Arg Ala Gly Ile Asn Val Arg Glu Gln Thr Ser  
 450 455 460  
 Asn Pro Asn Ser Leu Leu Asn Tyr Tyr Lys Arg Leu Ile His Leu Arg  
 465 470 475 480  
 Arg Glu Thr Pro Ala Leu Ile Ala Gly Asp Tyr Val Pro Leu His Gln  
 485 490 495  
 Thr Ser Lys Asp His Leu Ala Phe Leu Arg Lys Thr Asp Ser Gln Thr  
 500 505 510  
 Ile Leu Val Val Leu Asn Tyr Ser Pro Asn Lys Leu Glu Leu Asp Phe  
 515 520 525  
 Ser Arg Thr Val Glu Met Lys Gly Arg Pro Leu Ile Ala Ile Phe Ser  
 530 535 540  
 Ser Ala Asp Asp Arg Pro Gln Ala Ala Gln Ser Pro Lys Lys Val Ser  
 545 550 555 560  
 Val Gly Ala Tyr Gly Val Leu Leu Ala Glu Val Lys  
 565 570

&lt;210&gt; 13

&lt;211&gt; 1392

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;400&gt; 13

atggaagaag	agcctagggg	taaaggccta	aagatcgcat	ttataggtgc	tgggagcgcc	60
gtatggagct	ctaggataat	tgtagacctt	atcctcgcta	aaagcctgca	tggcgccaag	120
atttacctca	tgatatacaa	taaggagcgc	ctcgacctca	taacagggtt	cgccaagcga	180
tacgctgccg	agatgcacgc	tgatctagag	ttcatcccga	caatggatcg	tgtcgaagcg	240
ataagggatg	ccgacttcgt	cgtcaactca	gccatgtatg	gtggacacat	gtactatgag	300
cggatgaggg	aaatatgtga	gagacacggc	tattacaggg	gaataaatag	tgtcgagtgg	360
aacatggtca	gcgactacca	caccatatgg	ggctactacc	agttcaaact	agcactaagc	420
attgccaaag	acgtcgagga	ttacgctccc	gatgcatggc	ttatcaacgt	cgctaatacca	480
gtgttcgaat	tgacaacact	gatccagagg	cacgtcaaga	taaagatgat	agggctgtgc	540
catggatacc	acggcatata	taatgtcatg	aaagaactag	gcctaccacg	agaagaaaca	600
gagttcgagg	tactaggatt	caaccacgtc	atatggctta	caaagttcaa	gtaccagggg	660
gaagacgctt	acccgttact	agacaagtgg	atcgaggaga	aagcagagaa	gtactgggag	720
cattggagac	aaacacaggt	aaaccggttc	gacatcgact	tgctgccggc	agcgatagac	780
atgtacaaga	gatacggctc	cctcccggta	ggagacactg	tgctgtggag	gacgtggatg	840
taccactggg	atctcaagac	gaagcagaaa	tggtatggcc	cgacaggagg	accagactcc	900
gagataggct	ggatgatgta	tatagccttc	ctaagcatgc	agctccaaag	actatacgaa	960
gcactaacgg	atcagaagca	cccgttagca	gcacatgtac	cgccggagtg	gagcgggtgaa	1020
tccatagctc	caataatcga	tagcctcgcc	aacaatagga	ggggagaata	cgttatcaac	1080
acgttgaatc	tgggcagcat	acccgggata	ccggataatg	tggtgtgtcga	gatgccggcc	1140
cagatagatg	gtaaaggagt	gcaccgctac	atattcgagc	cactccccaa	gaagataaga	1200
gacctggtcc	tactgcctag	gatgacccgt	atggagatgg	cattgacagc	cttcctcgag	1260
ggaggccgtg	aagtactaga	ggactggcta	cacatggatc	cacgtaccaa	gagcactaga	1320
caggtacggg	agacaatcga	tgatctcctt	aacatgcccg	gtaacgagga	gatgaagaag	1380
catttcagct	aa					1392

&lt;210&gt; 14

&lt;211&gt; 463

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;221&gt; DOMAIN

&lt;222&gt; (21)...(446)

&lt;223&gt; Family 4 glycosyl hydrolase

&lt;400&gt; 14

Met	Glu	Glu	Glu	Pro	Arg	Gly	Lys	Gly	Leu	Lys	Ile	Ala	Phe	Ile	Gly
1				5				10						15	
Ala	Gly	Ser	Ala	Val	Trp	Ser	Ser	Arg	Ile	Ile	Val	Asp	Leu	Ile	Leu
			20					25					30		
Ala	Lys	Ser	Leu	His	Gly	Ala	Lys	Ile	Tyr	Leu	Met	Asp	Ile	Asn	Lys
			35				40					45			
Glu	Arg	Leu	Asp	Leu	Ile	Thr	Gly	Phe	Ala	Lys	Arg	Tyr	Ala	Ala	Glu
			50			55					60				
Met	His	Ala	Asp	Leu	Glu	Phe	Ile	Pro	Thr	Met	Asp	Arg	Val	Glu	Ala
65				70					75					80	
Ile	Arg	Asp	Ala	Asp	Phe	Val	Val	Asn	Ser	Ala	Met	Tyr	Gly	Gly	His
			85					90					95		
Met	Tyr	Tyr	Glu	Arg	Met	Arg	Glu	Ile	Cys	Glu	Arg	His	Gly	Tyr	Tyr
			100				105					110			
Arg	Gly	Ile	Asn	Ser	Val	Glu	Trp	Asn	Met	Val	Ser	Asp	Tyr	His	Thr
			115			120						125			
Ile	Trp	Gly	Tyr	Tyr	Gln	Phe	Lys	Leu	Ala	Leu	Ser	Ile	Ala	Lys	Asp
			130			135					140				
Val	Glu	Asp	Tyr	Ala	Pro	Asp	Ala	Trp	Leu	Ile	Asn	Val	Ala	Asn	Pro

145 Val Phe Glu Leu Thr Thr Leu Ile Gln Arg His Val Lys Ile Lys Met 160  
 165 Ile Gly Leu Cys His Gly Tyr His Gly Ile Tyr Asn Val Met Lys Glu 175  
 180 Leu Gly Leu Pro Arg Glu Glu Thr Glu Phe Glu Val Leu Gly Phe Asn 190  
 195 His Val Ile Trp Leu Thr Lys Phe Lys Tyr Gln Gly Glu Asp Ala Tyr 205  
 210 Pro Leu Leu Asp Lys Trp Ile Glu Glu Lys Ala Glu Lys Tyr Trp Glu 220  
 225 His Trp Arg Gln Thr Gln Val Asn Pro Phe Asp Ile Asp Leu Ser Pro 235  
 245 Ala Ala Ile Asp Met Tyr Lys Arg Tyr Gly Leu Leu Pro Val Gly Asp 255  
 260 Thr Val Arg Gly Gly Thr Trp Met Tyr His Trp Asp Leu Lys Thr Lys 270  
 275 Gln Lys Trp Tyr Gly Pro Thr Gly Gly Pro Asp Ser Glu Ile Gly Trp 285  
 290 Met Met Tyr Ile Ala Phe Leu Ser Met Gln Leu Gln Arg Leu Tyr Glu 300  
 305 Ala Leu Thr Asp Gln Lys His Pro Leu Ala Ala His Val Pro Pro Glu 315  
 325 Trp Ser Gly Glu Ser Ile Val Pro Ile Ile Asp Ser Leu Ala Asn Asn 335  
 340 Arg Arg Gly Glu Tyr Val Ile Asn Thr Leu Asn Leu Gly Ser Ile Pro 350  
 355 Gly Ile Pro Asp Asn Val Ala Val Glu Met Pro Ala Gln Ile Asp Gly 365  
 370 Lys Gly Val His Arg Tyr Ile Phe Glu Pro Leu Pro Lys Lys Ile Arg 380  
 385 Asp Leu Val Leu Leu Pro Arg Met Thr Arg Met Glu Met Ala Leu Thr 395  
 405 Ala Phe Leu Glu Gly Gly Arg Glu Val Leu Glu Asp Trp Leu His Met 415  
 420 Asp Pro Arg Thr Lys Ser Thr Arg Gln Val Arg Glu Thr Ile Asp Asp 430  
 435 Leu Leu Asn Met Pro Gly Asn Glu Glu Met Lys Lys His Phe Ser 445  
 450 455 460

<210> 15

<211> 1632

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 15

atgacgcaat	cctcctccgc	cgtttcccca	gcctcctcgg	gtcccgcctg	gtggcatggc	60
gcgtcgatct	accagattta	cccgcgcagt	tttgccgaca	gcaacgggtga	tggcgtgggc	120
gaccttgccg	ggattacttc	gcatctggag	catatcgcca	gcctgggtgt	ggaggcgatc	180
tggatcagcc	cgtttttcac	cagcccgatg	gccgattatg	gctatgacgt	ggcggattat	240
tgcatggtcg	atccgatctt	cgggactttg	gcggacttcg	atgcgctggg	cgagaaggcc	300
catgggctgg	gcctgaaggt	caccatcgac	atgggtcttg	cccataccag	cgacaggcat	360
ccgtggttcg	aacaatcgcg	ttcggcgcg	gagaatgacc	gcgcccactg	gtacgtctgg	420
gccgatccca	agccggacgg	cacgccgccc	aacaactggc	agtcggtgtt	tggcggcccc	480
gcctggacct	gggacgcgcg	gcgcgggcag	tactacatgc	accagttcct	gaaggagcag	540
ccgcaattga	acgcgcacaa	tcctgcgggtg	caggatgcgc	tgctcgatgc	cttgcgcttc	600

```

tggctggagc ggggggtgga cgggttccgg ctcgatgcgc tgaaccactc gatgttcgat 660
ccggcgctga ccgacaatct gcccgcgccc gaggatggca agatccgcac gcggcctttc 720
gattttcagt tgaaaatcaa cagccagaac catcccgctg tgacgctgtt catcgagcgg 780
atcgccgatg tttgcggcca gcatggcgcg gtcttcaccg tggccgaagt gggcggcgat 840
ggcgcggttc cgctgatgaa ggcgtaacag gcgggagcgc atcggtgtc ctcggcctac 900
agctttgatt tcctttatgc gccggccttg acgggagcgc tggtaggcaa tgctctggcc 960
cagtggacgg gcaagccggg ggccgatggg ctgagcgaag gctggcccag ctgggcgttt 1020
gagaaccatg atgcgcgcgc ccataatctcg cgctgggtgg gcgaggagca tcgcgcgcgc 1080
tttgcgggga tgagccttgt gctgctggcc tcgctgcgcg gtaacatgtt catgtatcag 1140
gggcaggaac tggcgctgga gcaggatgag atcccgcttc atctgctgaa agaccccgag 1200
gcgatcgcca actggcgcgt gacgttgagc cgcgacgggg tgcgacgccc gatgccaatgg 1260
gacagccagg ccttccatgc cggcttcacg agtggcgagc cctgggttggc cttgtcgccg 1320
gggaatatcg ccaaggcggg ggatgtgcag gaggccgatc cgcagagcca gctgcactgg 1380
gtgcggcggg ttctggccct gcgcgcgcgg cacaaggccc tgcgtctggg cgcgatggag 1440
catgtgcatg tgcagggcga tgtgctcagt ttcaccgcgc atgcgcgcgg cgaaaggggtg 1500
gagtgcgtgt tcaacctttc cgccaaaacc gtcgccaca aggcgcacaa gggtagagacg 1560
ctgctcaccg tcaatggcgc aaccggggcg gtgttgacgc catatggcgc tctctggacg 1620
aaactcgcat aa 1632

```

<210> 16  
 <211> 543  
 <212> PRT  
 <213> Unknown

<220>  
 <223> Obtained from an environmental sample

<221> DOMAIN  
 <222> (25)...(415)  
 <223> Alpha amylase, catalytic domain

```

<400> 16
Met Thr Gln Ser Ser Ser Ala Val Ser Pro Ala Ser Ser Gly Pro Ala
1      5      10      15
Trp Trp His Gly Ala Ser Ile Tyr Gln Ile Tyr Pro Arg Ser Phe Ala
20     25     30
Asp Ser Asn Gly Asp Gly Val Gly Asp Leu Ala Gly Ile Thr Ser His
35     40     45
Leu Glu His Ile Ala Ser Leu Gly Val Glu Ala Ile Trp Ile Ser Pro
50     55     60
Phe Phe Thr Ser Pro Met Ala Asp Tyr Gly Tyr Asp Val Ala Asp Tyr
65     70     75     80
Cys Asp Val Asp Pro Ile Phe Gly Thr Leu Ala Asp Phe Asp Ala Leu
85     90     95
Val Glu Lys Ala His Gly Leu Gly Leu Lys Val Thr Ile Asp Met Val
100    105    110
Phe Ala His Thr Ser Asp Arg His Pro Trp Phe Glu Gln Ser Arg Ser
115    120    125
Ala Arg Glu Asn Asp Arg Ala Asp Trp Tyr Val Trp Ala Asp Pro Lys
130    135    140
Pro Asp Gly Thr Pro Pro Asn Asn Trp Gln Ser Val Phe Gly Gly Pro
145    150    155    160
Ala Trp Thr Trp Asp Ala Arg Arg Gly Gln Tyr Tyr Met His Gln Phe
165    170    175
Leu Lys Glu Gln Pro Gln Leu Asn Ala His Asn Pro Ala Val Gln Asp
180    185    190
Ala Leu Leu Asp Ala Leu Arg Phe Trp Leu Glu Arg Gly Val Asp Gly
195    200    205
Phe Arg Leu Asp Ala Leu Asn His Ser Met Phe Asp Pro Ala Leu Thr

```



210  
 Asp Asn Leu Pro Ala Pro Glu Asp Gly Lys Ile Arg Thr Arg Pro Phe  
 225 230 235 240  
 Asp Phe Gln Leu Lys Ile Asn Ser Gln Asn His Pro Ala Val Thr Leu  
 245 250 255  
 Phe Ile Glu Arg Ile Ala Asp Val Cys Gly Gln His Gly Ala Val Phe  
 260 265 270  
 Thr Val Ala Glu Val Gly Gly Asp Gly Ala Val Pro Leu Met Lys Ala  
 275 280 285  
 Tyr Thr Ala Gly Glu His Arg Leu Ser Ser Ala Tyr Ser Phe Asp Phe  
 290 295 300  
 Leu Tyr Ala Pro Ala Leu Thr Gly Glu Leu Val Ala Asn Ala Leu Ala  
 305 310 315 320  
 Gln Trp Thr Gly Lys Pro Gly Ala Asp Gly Leu Ser Glu Gly Trp Pro  
 325 330 335  
 Ser Trp Ala Phe Glu Asn His Asp Ala Pro Arg His Ile Ser Arg Trp  
 340 345 350  
 Val Gly Glu Glu His Arg Ala Ala Phe Ala Arg Met Ser Leu Val Leu  
 355 360 365  
 Leu Ala Ser Leu Arg Gly Asn Met Phe Met Tyr Gln Gly Gln Glu Leu  
 370 375 380  
 Ala Leu Glu Gln Asp Glu Ile Pro Phe His Leu Leu Lys Asp Pro Glu  
 385 390 395 400  
 Ala Ile Ala Asn Trp Pro Leu Thr Leu Ser Arg Asp Gly Val Arg Thr  
 405 410 415  
 Pro Met Pro Trp Asp Ser Gln Ala Phe His Ala Gly Phe Thr Ser Gly  
 420 425 430  
 Glu Pro Trp Leu Pro Leu Ser Pro Gly Asn Ile Ala Lys Ala Val Asp  
 435 440 445  
 Val Gln Glu Ala Asp Pro Gln Ser Gln Leu His Trp Val Arg Arg Val  
 450 455 460  
 Leu Ala Leu Arg Ala Arg His Lys Ala Leu Arg Leu Gly Ala Met Glu  
 465 470 475 480  
 His Val His Val Gln Gly Asp Val Leu Ser Phe Thr Arg His Ala Arg  
 485 490 495  
 Gly Glu Arg Val Glu Cys Val Phe Asn Leu Ser Ala Lys Thr Val Ala  
 500 505 510  
 His Lys Ala His Lys Gly Glu Thr Leu Leu Thr Val Asn Gly Ala Thr  
 515 520 525  
 Gly Ala Val Leu Thr Pro Tyr Gly Ala Leu Trp Thr Lys Leu Ala  
 530 535 540

&lt;210&gt; 17

&lt;211&gt; 1422

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;400&gt; 17

atgaagtaca	agaaccagtt	aaaaccactc	atggaagaag	agcctcgggg	taaaggcctg	60
aagatcgcg	ttataggggc	tgggagcgcc	gtctggagct	ctaggataat	catagacctt	120
atcctcgcg	aaagcctgca	tggcgccaaa	gtctatctca	tggatattaa	taaggagcgg	180
ctcgacctca	taacaggttt	cgccaagcga	tacgctgccg	agatgcatgc	tgatctagag	240
ttatcccta	caatggatcg	catagaggcg	ataagggatg	cagacttcgt	cgtaactca	300
gccatgtacg	gtggccacat	gtactatgag	cgtatgagag	aaatatgtga	gaggcacggc	360
tattaccgag	gaataaacag	tggtgaatgg	aacatgggtca	gcgattacca	caccatatgg	420
ggttactacc	agttcaaact	agccatgagc	attgccaagg	atgtagagga	atacgcgccc	480

```

gacgcctggc ttatcaacgt cgccaaccct gtgttcgagc tgacaacact gatccagagg 540
cacgtcaaga taaagatgat agggctctgc catggctacc atggcatcta taatgtgatc 600
aaagaactcg ggctagacag ggatgagaca gagttcgagg tactcggtt caaccatgtg 660
atctggctca caaagttcaa gtacagggga gaggacgctt accccttact agataagtgg 720
atcgaggaga aagcagagaa gtactgggag cattgggagac aaacacaggt gaacccgttc 780
gacattgact tgtctcctgc agcgatagac atgtacaaga gatacgggtct acttccggta 840
ggagatactg tgcgtggagg cacgtggatg taccactggg atctcaagac gaagcagaaa 900
tggtatggac cgacaggagg accagactcc gagataggct ggatgatgta tatagccttc 960
ctaagcatgc agctccaaag attatacgaa gcactaatgg atcagaagca cccattagca 1020
gcacatatac cgccggagtg gagcgggtgaa tccatagtcc caataatcga tagcctcgcc 1080
aacaatagga ggggagaata cgttatcaac acgttgaacc ttggtagtat accggggata 1140
ccggatagtg tagctgtcga gatgccagcc cagatagatg gtaaaggagt gcaccgctac 1200
atattcgagc ccttcccaa gaagataaga gacctagtcc tactgcctag gatgaccggt 1260
atggagatgg cgttgacagc cttcctcgag ggaggccgag aagtactcga ggactggcta 1320
cacatggatc cacgtaccaa gagcactgga caggtagcgg agacaatcga tgatctcctg 1380
aacatgcccg gtaacgagga gatgaagaag catttcagct ag 1422

```

&lt;210&gt; 18

&lt;211&gt; 473

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;221&gt; DOMAIN

&lt;222&gt; (21)...(446)

&lt;223&gt; Family 4 glycosyl hydrolase

&lt;400&gt; 18

```

Met Lys Tyr Lys Asn Gln Leu Lys Pro Leu Met Glu Glu Glu Pro Arg
 1          5          10          15
Gly Lys Gly Leu Lys Ile Ala Phe Ile Gly Ala Gly Ser Ala Val Trp
 20          25          30
Ser Ser Arg Ile Ile Ile Asp Leu Ile Leu Ala Lys Ser Leu His Gly
 35          40          45
Ala Lys Val Tyr Leu Met Asp Ile Asn Lys Glu Arg Leu Asp Leu Ile
 50          55          60
Thr Gly Phe Ala Lys Arg Tyr Ala Ala Glu Met His Ala Asp Leu Glu
 65          70          75          80
Phe Ile Pro Thr Met Asp Arg Ile Glu Ala Ile Arg Asp Ala Asp Phe
 85          90          95
Val Val Asn Ser Ala Met Tyr Gly Gly His Met Tyr Tyr Glu Arg Met
100          105          110
Arg Glu Ile Cys Glu Arg His Gly Tyr Tyr Arg Gly Ile Asn Ser Val
115          120          125
Glu Trp Asn Met Val Ser Asp Tyr His Thr Ile Trp Gly Tyr Tyr Gln
130          135          140
Phe Lys Leu Ala Met Ser Ile Ala Lys Asp Val Glu Glu Tyr Ala Pro
145          150          155          160
Asp Ala Trp Leu Ile Asn Val Ala Asn Pro Val Phe Glu Leu Thr Thr
165          170          175
Leu Ile Gln Arg His Val Lys Ile Lys Met Ile Gly Leu Cys His Gly
180          185          190
Tyr His Gly Ile Tyr Asn Val Ile Lys Glu Leu Gly Leu Asp Arg Asp
195          200          205
Glu Thr Glu Phe Glu Val Leu Gly Phe Asn His Val Ile Trp Leu Thr
210          215          220
Lys Phe Lys Tyr Arg Gly Glu Asp Ala Tyr Pro Leu Leu Asp Lys Trp

```

19

```

ggaacgcctt atatctatca aggcgaagaa ataggtatga caaaccggaa gttttcctct 1140
attgatgaat atagagatgt ggagtcatta aatgtgtatg aaataaaacg tgcacaagga 1200
atggacgaaa atgaaatttt ggaaatttta aaacataaat caagagataa ttcccgta 1260
ccggtgcaat ggaacgataa gccgaatgca ggttttacaa aaggaaagcc atggattcat 1320
ccggccgata actaccgtaa aattaatgta gaaaaagcgt tagaggataa agattcaatc 1380
ttttattttt atcaaaagct tattgcactg cgcaagcagt acgagattat cacctatgga 1440
aactatgaat tgattcttgg agaagacgag cagattttcg cttatatccg aaatggagca 1500
gatgaaaagc tgctcgtgat aaacaatttc tacggcagcg agaaaatttt tgaactgcca 1560
gaaaatctaa cttttgaagg atatcatagt gaaatattgc tgtctaacta cgaagattca 1620
ccaaaggaat tcaagcgagt cttacttcgg ccgtatgaat caatcgtgta tcatttaaaa 1680
aaatag

```

&lt;210&gt; 20

&lt;211&gt; 561

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;221&gt; DOMAIN

&lt;222&gt; (13)...(419)

&lt;223&gt; Alpha amylase, catalytic domain

&lt;400&gt; 20

```

Met Gln Glu Pro Trp Trp Lys Lys Ser Val Val Tyr Gln Ile Tyr Pro
1          5          10          15
Lys Ser Phe Tyr Asp Thr Thr Gly Asn Gly Val Gly Asp Ile Ala Gly
20          25          30
Ile Ile Glu Lys Leu Asp Tyr Leu Lys Lys Leu Gly Val Asp Val Leu
35          40          45
Trp Leu Thr Pro Ile Tyr Lys Ser Pro Gln Arg Asp Asn Gly Tyr Asp
50          55          60
Ile Ser Asp Tyr Phe Val Ile Gln Glu Glu Tyr Gly Thr Met Glu Asp
65          70          75          80
Phe Asp Leu Leu Val Thr Glu Ala His Lys Arg Gly Leu Lys Val Ile
85          90          95
Met Asp Ile Val Val Asn His Thr Ser Thr Glu His Glu Trp Phe Gln
100         105         110
Glu Ala Lys Lys Ser Lys Asp Asn Pro Tyr Arg Asp Phe Tyr Ile Trp
115         120         125
Lys Asp Gln Lys Glu Asp Gly Ser Ala Pro Thr Asn Trp Val Ser Lys
130         135         140
Phe Gly Gly Ser Ala Trp Glu His Asp Asn Leu Thr Glu Gln Ser Tyr
145         150         155         160
Leu His Leu Phe Asp Val Thr Gln Ala Asp Leu Asn Trp Glu Asn Glu
165         170         175
Arg Val Arg Arg Ser Val Tyr Asp Met Met Thr Phe Trp Phe Glu Lys
180         185         190
Gly Val Asp Gly Phe Arg Leu Asp Val Ile Asn Leu Ile Ser Lys Asp
195         200         205
Gln Arg Phe Leu Asp Asp Asp Gly Ser Val Ala Pro Gly Asp Gly Arg
210         215         220
Lys Phe Tyr Thr Asp Gly Pro Arg Val His Glu Tyr Met Arg Glu Met
225         230         235         240
Asn Gln Glu Val Phe Ser Lys Tyr Asp Ser Met Thr Val Gly Glu Met
245         250         255
Ser Ser Thr Thr Val Asp His Cys Ile Gln Tyr Ser His Pro Asp Arg
260         265         270

```

Arg Glu Leu Ser Met Thr Phe Asn Phe His His Leu Lys Val Asp Tyr  
 275 280 285  
 Pro Asn Gly Glu Lys Trp Ala Leu Ala Asp Phe Asp Phe Ile Lys Leu  
 290 295 300  
 Lys Glu Ile Leu Ser Thr Trp Gln Thr Glu Met Asn Lys Gly Gly Gly  
 305 310 315 320  
 Trp Asn Ala Leu Phe Trp Cys Asn His Asp Gln Pro Arg Val Val Ser  
 325 330 335  
 Arg Tyr Gly Asp Asp Glu Leu Tyr His Asn Lys Ser Ala Lys Met Leu  
 340 345 350  
 Ala Thr Thr Ile His Leu Met Gln Gly Thr Pro Tyr Ile Tyr Gln Gly  
 355 360 365  
 Glu Glu Ile Gly Met Thr Asn Pro Lys Phe Ser Ser Ile Asp Glu Tyr  
 370 375 380  
 Arg Asp Val Glu Ser Leu Asn Val Tyr Glu Ile Lys Arg Ala Gln Gly  
 385 390 395 400  
 Met Asp Glu Asn Glu Ile Leu Glu Ile Leu Lys His Lys Ser Arg Asp  
 405 410 415  
 Asn Ser Arg Thr Pro Val Gln Trp Asn Asp Lys Pro Asn Ala Gly Phe  
 420 425 430  
 Thr Lys Gly Lys Pro Trp Ile His Pro Ala Asp Asn Tyr Arg Lys Ile  
 435 440 445  
 Asn Val Glu Lys Ala Leu Glu Asp Lys Asp Ser Ile Phe Tyr Phe Tyr  
 450 455 460  
 Gln Lys Leu Ile Ala Leu Arg Lys Gln Tyr Glu Ile Ile Thr Tyr Gly  
 465 470 475 480  
 Asn Tyr Glu Leu Ile Leu Gly Glu Asp Glu Gln Ile Phe Ala Tyr Ile  
 485 490 495  
 Arg Asn Gly Ala Asp Glu Lys Leu Leu Val Ile Asn Asn Phe Tyr Gly  
 500 505 510  
 Ser Glu Lys Ile Phe Glu Leu Pro Glu Asn Leu Thr Phe Glu Gly Tyr  
 515 520 525  
 His Ser Glu Ile Leu Leu Ser Asn Tyr Glu Asp Ser Pro Lys Glu Phe  
 530 535 540  
 Lys Arg Val Leu Leu Arg Pro Tyr Glu Ser Ile Val Tyr His Leu Lys  
 545 550 555 560  
 Lys

&lt;210&gt; 21

&lt;211&gt; 1443

&lt;212&gt; DNA

&lt;213&gt; Eukaryote

&lt;400&gt; 21

atgccagctg	tgaagatcgg	tatcatcggg	gcaggaagtg	ccgtttttctc	actgaggctc	60
gtgagcgacc	tttgcaaaac	tcccgggtctt	tccggaagta	ccgttaccct	catggacatc	120
gacgaagaaa	ggctagatgc	ggttctgacc	atagcaaaaa	agtacgttga	agaagtggga	180
gccgacctga	agtttgaaaa	gacaacaagc	gtagacgaag	ccatcgctga	tgccgatttt	240
gtgataaaca	cagcgatggt	gggtggccac	acctatcttg	aaaagggtcag	aaggatcagc	300
gaaaagtacg	tatactacag	aggaatagac	gcgcaggagt	tcaacatggt	ctccgactac	360
tacacgtttt	caaactacaa	ccagctcaag	tacttcgtgg	atatcgcaag	gaaaatagag	420
agactctctc	caaaggcgtg	gtatctgcag	gcggcaaadc	ctgtctttga	aggaacaacc	480
cttgtgacaa	gaacggttcc	cataaaggca	gtgggattct	gccatggaca	ctacggcgtg	540
atggagatcg	tagagaaact	ggggctggaa	gaaaacaggg	tagactggca	ggttgccggt	600
gtgaaccacg	gaatatggtt	gaacaggttc	agatacaacg	gcgaggatgc	gtacccactt	660
ctagatagat	ggatcagtga	aaaatcgaaa	gactggaaac	cagagaaccc	cttcaacgac	720
cagctctctc	ccgctgcgat	agacatgtac	agattctacg	gtgtgatgcc	catcggtgac	780
accgtgagaa	actcttcgtg	gaggtaccac	agggatcttg	agaccaagaa	gaaatggtac	840

```

ggtgaaccct ggggaggagc agattctgaa ataggctgga aatggtacca ggacacactt      900
ggaaaaggtaa cggagatcac aaagaagggtg gcaaagttca tcaaagaaaa tccgtccgcg      960
aggctctccg accttggaag tgttcttgga aaggatcttt cggaaaaaca gttcgtcctt     1020
gaagtggaaa agattctcga tcctgaaaaa aagagcgggtg aacagcacat tcctttcatc     1080
gatgccctgc tcaacgacaa cagaagtaga tttgtgataa acattccaaa caaaggatc      1140
atacagggtg tgcagcagca tgtggttgtt gaagttccag cagttgtgga cagggacggt      1200
atccatccag aaaagatcgc cccaccactt ccagaacgtg tggatgaagta ttatctgaga      1260
ccgaggatca tgagaatgga gatggcactc gaagcgttcc tcacgggtga tataaggatc      1320
ataaaagaag ttctctacag agacccaagg acaaagagtg acgaacaggt agagaagggt      1380
atagaagaaa tcctttctct tccagaaaac gaagagatga gaaagaatta tctgaaaaaa      1440
taa                                                                    1443

```

<210> 22  
 <211> 480  
 <212> PRT  
 <213> Eukaryote

<220>  
 <221> DOMAIN  
 <222> (5)...(451)  
 <223> Family 4 glycosyl hydrolase

```

<400> 22
Met Pro Ala Val Lys Ile Gly Ile Ile Gly Ala Gly Ser Ala Val Phe
1          5          10          15
Ser Leu Arg Leu Val Ser Asp Leu Cys Lys Thr Pro Gly Leu Ser Gly
20          25          30
Ser Thr Val Thr Leu Met Asp Ile Asp Glu Glu Arg Leu Asp Ala Val
35          40          45
Leu Thr Ile Ala Lys Lys Tyr Val Glu Glu Val Gly Ala Asp Leu Lys
50          55          60
Phe Glu Lys Thr Thr Ser Val Asp Glu Ala Ile Ala Asp Ala Asp Phe
65          70          75          80
Val Ile Asn Thr Ala Met Val Gly Gly His Thr Tyr Leu Glu Lys Val
85          90          95
Arg Arg Ile Ser Glu Lys Tyr Gly Tyr Arg Gly Ile Asp Ala Gln
100          105          110
Glu Phe Asn Met Val Ser Asp Tyr Tyr Thr Phe Ser Asn Tyr Asn Gln
115          120          125
Leu Lys Tyr Phe Val Asp Ile Ala Arg Lys Ile Glu Arg Leu Ser Pro
130          135          140
Lys Ala Trp Tyr Leu Gln Ala Ala Asn Pro Val Phe Glu Gly Thr Thr
145          150          155          160
Leu Val Thr Arg Thr Val Pro Ile Lys Ala Val Gly Phe Cys His Gly
165          170          175
His Tyr Gly Val Met Glu Ile Val Glu Lys Leu Gly Leu Glu Glu Asn
180          185          190
Arg Val Asp Trp Gln Val Ala Gly Val Asn His Gly Ile Trp Leu Asn
195          200          205
Arg Phe Arg Tyr Asn Gly Glu Asp Ala Tyr Pro Leu Leu Asp Arg Trp
210          215          220
Ile Ser Glu Lys Ser Lys Asp Trp Lys Pro Glu Asn Pro Phe Asn Asp
225          230          235          240
Gln Leu Ser Pro Ala Ala Ile Asp Met Tyr Arg Phe Tyr Gly Val Met
245          250          255
Pro Ile Gly Asp Thr Val Arg Asn Ser Ser Trp Arg Tyr His Arg Asp
260          265          270
Leu Glu Thr Lys Lys Lys Trp Tyr Gly Glu Pro Trp Gly Gly Ala Asp
275          280          285

```

Ser Glu Ile Gly Trp Lys Trp Tyr Gln Asp Thr Leu Gly Lys Val Thr  
 290 295 300  
 Glu Ile Thr Lys Lys Val Ala Lys Phe Ile Lys Glu Asn Pro Ser Ala  
 305 310 315 320  
 Arg Leu Ser Asp Leu Gly Ser Val Leu Gly Lys Asp Leu Ser Glu Lys  
 325 330 335  
 Gln Phe Val Leu Glu Val Glu Lys Ile Leu Asp Pro Glu Lys Lys Ser  
 340 345 350  
 Gly Glu Gln His Ile Pro Phe Ile Asp Ala Leu Leu Asn Asp Asn Arg  
 355 360 365  
 Ser Arg Phe Val Ile Asn Ile Pro Asn Lys Gly Ile Ile Gln Gly Ile  
 370 375 380  
 Asp Asp Asp Val Val Val Glu Val Pro Ala Val Val Asp Arg Asp Gly  
 385 390 395 400  
 Ile His Pro Glu Lys Ile Ala Pro Pro Leu Pro Glu Arg Val Val Lys  
 405 410 415  
 Tyr Tyr Leu Arg Pro Arg Ile Met Arg Met Glu Met Ala Leu Glu Ala  
 420 425 430  
 Phe Leu Thr Gly Asp Ile Arg Ile Ile Lys Glu Val Leu Tyr Arg Asp  
 435 440 445  
 Pro Arg Thr Lys Ser Asp Glu Gln Val Glu Lys Val Ile Glu Glu Ile  
 450 455 460  
 Leu Ser Leu Pro Glu Asn Glu Glu Met Arg Lys Asn Tyr Leu Lys Lys  
 465 470 475 480

&lt;210&gt; 23

&lt;211&gt; 1686

&lt;212&gt; DNA

&lt;213&gt; Unknown

&lt;220&gt;

&lt;223&gt; Obtained from an environmental sample

&lt;400&gt; 23

atgcaagaac	cttggtggaa	aaaatcggtt	gtttatcaaa	tttatccaaa	aagttttttat	60
gatacgactg	gaaacggcgt	gggtgatata	gctggcatta	ttgaaaagct	agattatttg	120
aaaaagctcg	gggtagatgt	tctgtggcta	acgccaattt	ataaatcacc	gcagcgggat	180
aatggatatg	atataagtga	ttattttggt	attcaagaag	aatacggaac	aatggaggat	240
tttgatcttt	tagtaacaga	agcgcataag	cggggtctta	aagtcatcat	ggatattgtc	300
gttaatcata	catcaactga	acatgaatgg	tttcaagaag	ctaaaaaatc	gaaagataac	360
ccgtaccgag	atttttatat	ttggaaagat	caaaaagaag	atggaagtgc	tccgacgaat	420
tgggtttcaa	aatttggggg	atccgcgtgg	gagcatgata	acctcacaga	acaatcgtat	480
ctgcatttgt	ttgatgttac	gcaagcggat	ttgaactggg	aaaacgagcg	tgtgcgccgc	540
agcgtgtatg	atatgatgac	gttttggttt	gaaaaaggag	tagatggatt	tcgtctggac	600
gttattaatt	taattttcaa	agatcagcgt	tttttagatg	atgacggttc	cgttgcacca	660
ggagatggcc	ggaaattcta	caccgacggg	cctcgcgtgc	acgaatatat	gcgggaaatg	720
aatcaagaag	ttttttcaaa	atatgatagt	atgaccgttg	gagaaatgtc	gtccacaacc	780
gttgaccact	gtattcagta	ctctcatccg	gaccgacgcg	agcttagtat	gacgtttaat	840
tttcatcatt	tgaaagtcga	ttaccggaat	ggagaaaagt	gggccctagc	agattttgat	900
tttattaaat	taaaagagat	tctatcaact	tggcaaacgg	aaatgaataa	aggtggggga	960
tggaatgcat	tattttggtg	caaccatgat	cagcctcgcg	ttgtttcacg	ctatggagac	1020
gacgaactct	atcataataa	atctgctaaa	atgctcgcta	caacgattca	tttgatgcag	1080
ggaacgcctt	atattttatca	aggcgaagaa	ataggtatga	caaaccggaa	gttttcctct	1140
attgatgaat	atagagatgt	ggagtcatta	aatgtgtatg	aaataaaaacg	tgcacaggga	1200
atggacgaaa	atgaaathtt	ggaaatttta	aaacataaat	caagagataa	ttcccgtaca	1260
ccggtgcaat	ggaacgataa	gccaaatgca	ggttttacaa	aaggaaagcc	atggattcat	1320
ccggccgata	actaccgtaa	aattaatgta	gaaaaagcgt	tagaggataa	agattcaatc	1380
ttttattttt	atcaaaaagct	tattgcactg	cgcaagcagt	acgagattat	cacctatgga	1440
aactatgaat	tgattcttgg	agaagacgag	cagattttcg	cttatatccg	aaatggagca	1500

```

gatgaaaagc tgctcgtgat aaacaatttc tacggcagcg aaacagcctt tgaactgcca 1560
gaaaatctaa cttttgaagg atatcatagt gaaatattgc tgtctaacta cgaagattca 1620
ccaaaggaat tcaagcgagt cttacttcgg ccgtatgaat caatcgtgta tcatttaaaa 1680
aatag 1686

```

<210> 24  
 <211> 561  
 <212> PRT  
 <213> Unknown

<220>  
 <223> Obtained from an environmental sample

<221> DOMAIN  
 <222> (13)...(419)  
 <223> Alpha amylase, catalytic domain

<400> 24  
 Met Gln Glu Pro Trp Trp Lys Lys Ser Val Val Tyr Gln Ile Tyr Pro  
 1 5 10 15  
 Lys Ser Phe Tyr Asp Thr Thr Gly Asn Gly Val Gly Asp Ile Ala Gly  
 20 25 30  
 Ile Ile Glu Lys Leu Asp Tyr Leu Lys Lys Leu Gly Val Asp Val Leu  
 35 40 45  
 Trp Leu Thr Pro Ile Tyr Lys Ser Pro Gln Arg Asp Asn Gly Tyr Asp  
 50 55 60  
 Ile Ser Asp Tyr Phe Val Ile Gln Glu Glu Tyr Gly Thr Met Glu Asp  
 65 70 75 80  
 Phe Asp Leu Leu Val Thr Glu Ala His Lys Arg Gly Leu Lys Val Ile  
 85 90 95  
 Met Asp Ile Val Val Asn His Thr Ser Thr Glu His Glu Trp Phe Gln  
 100 105 110  
 Glu Ala Lys Lys Ser Lys Asp Asn Pro Tyr Arg Asp Phe Tyr Ile Trp  
 115 120 125  
 Lys Asp Gln Lys Glu Asp Gly Ser Ala Pro Thr Asn Trp Val Ser Lys  
 130 135 140  
 Phe Gly Gly Ser Ala Trp Glu His Asp Asn Leu Thr Glu Gln Ser Tyr  
 145 150 155 160  
 Leu His Leu Phe Asp Val Thr Gln Ala Asp Leu Asn Trp Glu Asn Glu  
 165 170 175  
 Arg Val Arg Arg Ser Val Tyr Asp Met Met Thr Phe Trp Phe Glu Lys  
 180 185 190  
 Gly Val Asp Gly Phe Arg Leu Asp Val Ile Asn Leu Ile Ser Lys Asp  
 195 200 205  
 Gln Arg Phe Leu Asp Asp Asp Gly Ser Val Ala Pro Gly Asp Gly Arg  
 210 215 220  
 Lys Phe Tyr Thr Asp Gly Pro Arg Val His Glu Tyr Met Arg Glu Met  
 225 230 235 240  
 Asn Gln Glu Val Phe Ser Lys Tyr Asp Ser Met Thr Val Gly Glu Met  
 245 250 255  
 Ser Ser Thr Thr Val Asp His Cys Ile Gln Tyr Ser His Pro Asp Arg  
 260 265 270  
 Arg Glu Leu Ser Met Thr Phe Asn Phe His His Leu Lys Val Asp Tyr  
 275 280 285  
 Pro Asn Gly Glu Lys Trp Ala Leu Ala Asp Phe Asp Phe Ile Lys Leu  
 290 295 300  
 Lys Glu Ile Leu Ser Thr Trp Gln Thr Glu Met Asn Lys Gly Gly Gly  
 305 310 315 320  
 Trp Asn Ala Leu Phe Trp Cys Asn His Asp Gln Pro Arg Val Val Ser



[illegible]